



Left Inverted Terminal Repeat: 1-103

Encapsidation Signal (Ψ): 183-331

HPRT Introns: 365-10547

Right Inverted Terminal Repeat:
10561-10663

pBR322 ori: 10867-11534

Kanamycin Resistance Gene: 12343-
13134

FIG. 1

SEQ ID NO: 1 pShuttle Sequence

CATCATCAATAATATACCTTATTTTGGATTGAAGCCAATATGATAATGAGG
GGGTGGAGTTTGTGACGTGGCGCGGGGCGTGGGAACGGGGCGGGTGACG
TAGTAGTGTGGCGGAAGTGTGATGTTGCAAGTGTGGCGGAACACATGTAA
GCGACGGATGTGGCAAAAGTGACGTTTTTGGTGTGCGCCGGTGTACACAG
GAAGTGACAATTTTCGCGCGGTTTTAGGCGGATGTTGTAGTAAATTTGGGC
GTAAACCGAGTAAGATTTGGCCATTTTCGCGGGGAAAACCTGAATAAGAGGAA
GTGAAATCTGAATAATTTTGTGTTACTCATAGCGCGTAATACTGGTACCGC
GGCCGCTCGAGTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCTG
ATGGCTCTCAAAATTCCTGCCTCCTTTAGGGATAAAAGACTTTAAGACTTT
TTAACAAAAAAGAAAAAGAAAAAATTCCTGCCTCCTGGTGTACACA
CACAGAAGGGTTCCTCCCCTTGAATGTGACCAGGATCTGTGAAAATAAC
GGGATAGCCGCTCCTGTGATTAGGTTATGTGGTAGACTAGAGCAAGATTC
TCCTGCTGGTTTTGAAGAAGTCAGCTGCCATGTTGTGAGACTGTCATGGGC
TAGGGCATGAGCCTTTAAATATCTGGGAGCAACCCCTGGCCAGCAGCCAG
TGAGAAAACGGGCCCTCAGTCCTACAATCACAAGGAACTAAATTCTGCCA
ACAACCTGAAGGAACTTTGAAGAGGATCATGAGTCCCTTGATTCAGCTTG
ATGAGCCCCTGAGCAGAGGATACAGCTAACTTGTACTAGGGAAGTATAAA
AAACATGCATGGGAATGATATATATCAACTTTAAGGATAATTGTCATACTT
CTGGGAATGAAGGGAAAGAAATGGGGCTTTAGTTGTATTATGATCTTTAA
TTTCTCAAAAAAATAAGATCAGAAGCAAATATGGCAAAATGTTAATACT
TTTGTGGGTACGTAGGTATTCAGCATACCCTTTTTTCTGAGTTCAAAATAT
TTTATAATTAAAATGAAATGCAGGCCAGGCACAGTGGCTCATGCCTATAA
TACCAGCACTTTGCGAGGCCGAGGTGGGAGGATGGCTTGAGGCCAGACCA
GCCTGGCCAACATGGCAAAACCCCATCTCTACTTAAAAAAAAAAAACTA
TATATATATATATGTGTGTGTGTGTATATATATATATGTATATATATTTA
TATATGTGTGTATATATATATATGTATATATATTTATATATGTGTGTGTATA
TATATATATACACACACACACATATATACATACATACATACACACACACA
CACACACAATTAGCCAGGCATGGTGGCGCACACCTGTAGTCCCAGCTACT
TGGGAGGCTGAGACATGAGAATTGCTTGAACCTGGGAGGCAGAGTAGTTA
GTGAGCTGAGATCATACCACTGCACTCCAGCCTGGTGACAGAGTGAGACT
CTGTCTTAAAAAAAAATAAAAAATTAATAATTAATGCAAAAGGTCCAAGTGA
ATTGAAGAGGAAAGGGGTATCAAGGAAGGTTTTGTGGAGGTGACGTTTGA
GCTGGGTCTTAAATGACTTAAACATGGGATAAGAAGGGAGGGAATAAGG
ACATTTCAAGTACGAGAAATAAGGAGCAAACAGTGGAAACAACCTAACG
TCTGTCAACCAGTGAATGGATAACAAAAATGTAATTCAGATGGTATCCAA
CTTACGATGGTTCAACATGAGATTTTTCTGACTTTAGGATAGATTTATCAA
AGTAGTAAATCCATTTTCAACTTATGATATTTTCAACTTCAGATGGGTTTA
TCAGGACACAGTTGAGGAACACCTGTCTATCCATACAATTTGGCAATAAA
AAGGAAATGAGTGCAGATATACTCCACAACATGAATGAACCTTGAAAACA
TTAAGTGAGAGAAGCCAGATACAAAAGGCCACATATTGTATGATTCTATT
TATACAAAATGTCCAGAATAGGCAAATCTTATAGACAGCAAGTAGGTAGA
TGATCAGTTTGCTAGGTGCTGGGGGAAGGGGAAATGGGGAGTGATGGCTA
AGGGGATTGGGTTTCTTTGTGGGGCAATGAAAATGTTTTAAATTTAGCG
TGATAATGATTGCACAATGCTGCATATATATATAATCTATAGATTATATAT
ATATAAAGAGAGGCTGTTAGACAGTGATAAGTGATATATATATATATATA
CATAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGTTAGTGATAAGTGATC
AGGAAAATAAAAGTATTGAGGAGGAATACGAAGTTGACGGTGTGAAAAC
ATGAGATTTTATATAGGATGGCCAGGGAAGGCCTTAATGAGAAAGTGACT

FIG. 2

TATGAGTAAAAACAAGGGATCCTAAACCTTAGCATGCATCAGAATCACTC
GGAAACTTGTTAAAGCATAGCTTGCTGGGCCTCATCACAGATATTTTGATT
CGGTAGGTTCTTGTCTGATATTAATACTTTTGGTCTAGGGAACCACATTTT
GAGAACCCTGAGCTAAAGGAAGTAAAGGTTTCCCTTAGTTTACTAGCTG
GTAACCCTAGGAACTGCTTAGCCTCTCGGTGCTAAGATACAAAATACTTT
AGCACATAATAACACATGGAAAATAGTCTATAAAATTATAAATATTATTTT
TATGTACCAAATATTACATAAGACAAAATCTAAGCAAGATATATATATAT
ATACATAAAATATAAGATATATATGTATATATTATATATAGATAAATAGA
GAGAGAGAGTTATGTTTAGAAAGAAAATACTTCAAACATAAAAAAGAGA
GGTAGGAAGTATACCATTCCATTATTGGTAAAAACAAATTACTAAGTAGT
CTTTACAAAAAACCAATCTCACTCCTTTAGAACACAAGCCCACCATTAAA
ACTGATGCAGAGGAATTTCTCTCCCTGGCTTACCTTTAGGATGGTGCATAC
TAAGTTAGAAAAGTCATAAATGTTATATTAAGTAAATGTGAACCTTACT
TCCACAATCAAGACATTCTAGAAGAAAAAGAGAAATGAAAATCAGTACA
ATGAATAAAACGGTATTTCCAATTATAAGTCAAATCACATCATAACAACC
CTAAGGAATTATCCAAACTCTTGTTTTTAGATGCTTTATTATATCAAACCTCT
CCTTTAAACAAGTGGCCCATCTGCTGGGATTTGGAAGCCTGTAATACTGA
AATTTTCATCATAATGGAAATTTTAAAAACAGAATTTGACCCACCTGTTTT
TAAACACTTTTCATTACTTAACAAGAGGTCTAATCTTGGGCAAGTCTTGAA
ATTTCTCTGGCCTTAGTTTCCCATGTGTTAAATGAAACTTGAAGCAGTTGG
TCTCTTATAGTCTCCTGACTCTAACATTCTAAGAATTATATTTGTACAATA
ACTCAAAAATCACATAATTTAATTTACCATATGGACTCCAAAATATATTTT
CTCATTAGGCTAAACTTGATCTGCATTTTCTGGATGTGTCCATATTCTTGG
ACTACACTAAACATGATACCAATGCTTCCTCTCACCATAAACCCCTCACTT
CGCTTTCTACATTTAAGAATTTTATAGCTGGAAGAGTCCTTAACAGAAAAT
ACCATCTAATAATTACCCCTCAAAATCGAGAAAGTCCTATCTGTTCTTATG
CTAGTTATAAGAATGAGGCAGCATTTCACATAATGGTTATAAACACTGCC
ACAAGAAGATTCATGATGTGTTGTTTATCTGTAGCTCTCATCATACTCTGT
CATATAACTATAGCATTAAAGATTTTAAATGTTCTATATATTCTTCTAAGACA
GTGTTTACCAGAGTAAGGCACAAAAGATCCACTGGTTTGCAAGAAAGATT
AGAACTTTTAAATTTTTTACCTCACCTTGTTTAAATCTATATTTTTGTATGTA
TTTTGTAACATATATATTATTATTACCATAAATCATATATAATTTAAAATG
CATATATTAGGGGTAAATGCTCAGGAACTTTTTATAAATTGGGCATGCA
AATACAAGTTTGAAGACTCACTGTTCTAGGTATTAAAAGTAAAGTTATAA
CCAAGTAAAGCTTCCACCTTTTCATGTCTCAAAGCAGTTTATTGTTGGAGG
TAAGATCTCTTAGAAGCCTAAACAGGTCCAAGTACAGAAATGAAGTAAGGC
TAGCCCATAACTTGTGGCAAGCAATTCATACTATTTCTCTCATGCTGAGCT
CTCCTCAGTGAAGCAGCTACTATAGACAACTGCAGCCTATTGGTAGCCTAT
TTTACAGGCAGGAAAAAAATTACTTTTTATTCAAAGTGGAACCTCAGGACA
TGGGGAGAAAAATGAATACAAAAAATAGGGTCAATCCAAAGGCACACAGC
AAATGAGTAACACAGTTATGTTTTTTTCCCATTTGTATGAGGTCCCAGTAA
ATTCTAAGTAACTGCAAATTTAATAATACTAAAAAAGCCATGCAATT
GTTCAAATGAATCCCAGCATGGTACAAGGAGTACAGACACTAGAGTCTAA
AAAACAAAAGAATGCCATTATTGAGTTTTTGAATTATATCAAGTAGTTAC
ATCTCTACTTAATAAATGAGAAAAACGAGGATAAGAGGCCATTTGATAAA
ATGAAAATAGCCAAGAAGTGGTATTAGAGACTTGAATACAGGTATTCGGG
TCCAAAGTTCATCTGCTCAAATACTAACTGGGGAAAAGAGGGAAAAATAT
TTATATACATATATATCTGCACACAAAAATACCCCCAAAAGACAAAATGA
GGCCAGGCAGGGTGGCTCACACCCGTAATCCCGGTACTTTGGGAGGCTGA
GGCAGGTGGATACCTGAGATCAGGAGTTGGAGATCAGCCTGGTCAACATG

FIG. 2a

GTGAAACCCTGTCTCTACTAAAGATAAAAAAATTAGCCAGGCATGGTGGC
 GTGCGCCTGTAATCCCAGCTACTTGGGAGTCTGAGGCAGGAGAATCACTT
 GAACTGGGAAGGGGAGGTTGCAGTGAGCCAAGATCGTACTACTGCACTCC
 AGCCTGGGCAGCAGAGTGAGACTCCATCACAAAAATAAATAAATAAATA
 AAATACAATGAAACAGAAAGTTCAAATAATCCCATAATCTTACCACCAAG
 AAATAACTTTTCACTCGTTATACTTATTGATTTTTTCCATAATAAATGTACTTT
 ACTGTGACTATCATGAAAAGAAAGTTATTTTAGAAACAGAGAACTGTTTC
 AGATCAAATCTATGTAGTAGAACAGAGCCATTAGGTGGGAAAGACGAGA
 TCAAACATAATCTCAGAAGGCCTAAAAGGCTAGGTCCATTCCAGCACTAA
 AAACCTGACCAGACAAGTAATGGCTTCAACAGCTTCTAAATATGGACAAAG
 CATGCTGAAAGGGAAGGACAGGTCTAACAGTGGTATATGAAATGAACAG
 GAGGGGCAAAGCTCATTTCTCCTCTGAAGTTTTCCAAAGATGCTGAGGAG
 GACATTAGTTTGACATGACCCTGATATGGGACAAGATAATTTTACAGAAG
 TTTTACATGTTAAAGTTTTCTTATAGATACTCATTCAAGTAAGCAATGAAC
 ACTAAAATCTAAAGAAAGAAAAGAGCTTTAGAGTCAGGTCTGTATTCAA
 TTCAAGCTCTACCACTTACTGGTTCTGTGACTTTGGGCAAGTCTTTTAACCT
 TATTAAGTCTTAATTTCTGATTTGTAAAATGGGGATATCGTCTCCCTCAC
 AGGATTGTTGTGAAACTTTTATGAGATTAATGCCTTTATATTTGGCATAGT
 GTAAGTAAACAATAACTGGCAGCTTCAAAAAAAAAAAGCAGTAGCATTCC
 ATCATTTATTATTGGTTACTCTCAAAAAGTTTTTCAATGTACTAGAAGATA
 AATATTCAAATACCTTAATATCTCCATTATTTTCAGGTAAACAGCATGCTC
 CTGAACAACCAATGGGTCAACAAATAAATTAAAAGGGAAATCTAAAAAC
 ATCTTGATATTAACTACATGGAAGCACAATATACCAAACCAATGGTTC
 AACTAGGAGAATTTTAAGGTACAAGAAAACCTTTGAGATTTCTTAAAA
 TAATAGTATGTCTGAATTTATTGAGTGATTACCAGAACTGTTGTAAGAG
 CTCTACTTGCATTATAGCACTTAATCCTCTTAACTCTATGGCTGCTATTATC
 AACCTCACCTAATCACATATGGGACACAGAGAGGTAAAGTAACCTTGCCC
 AAGGTCAGAGTTAGGAAGTACTAAGCCATGCTTTGAATCAGTTGTCAGGC
 TCCGGAACCTCACACTTTCAGCCACTACATAATACTGCTTTGCTATCTTTTA
 GGAAACTATGTGAGTCTACCTCACATAGACTCACATAGGTTTGTTTTTTTT
 TTTTTTTTAAAGGCTATCTTTTCCCCCATCAATGTTTTTTGAAGGATCCCAA
 ATTAGAGTCCCACAGAGGCAGACAGCAGTACTTGACAATATGGACATTTA
 AGGTAAATGTTGGATTCTACTGTCTTTTTACTACATGACCTAGGGAACGAT
 AATTAACCTAGACTGCTTCCAAGGGTTAAATAACCCATTTAGTTATACTAT
 GTAAATTATCTCTTAGTGATTGATTGAAAGCACACTGTTACTAATTGACTC
 GGTATGAAGTGCTTTTTTTTCTTCCCTTTCAAGATACATACCTTTCCAGTTA
 AAGTTGAGAGATCATCTCCACCAATTACTTTTATGTCCCCTGTTGACTGGT
 CATTCTAGTTAAAAAAAAAAAAAAAACTATATATATATATATCTACACACAC
 ATATGTATATGTATATCCTTATGTACACACACAACTTCAAATTAATGAG
 AACTAGAAGATTTGAGAAGTTAGCTAGCTAATATCCATAGCATTATGATA
 TTCTAAATGATATGAATTATAAGAATTAGGTTTCCTGAAATGAATGACTAG
 AAAACTTTCAAGTAGAGATTAGTAAAAATTAAAAAGTCCTAATCGGCCAT
 TACTGATTTGATGTTTTTAAGAGTCCTAAAAAATGGGTTACATCCATTTTT
 AAGTGGGTTAGTATTATAACAGCCACCCATCTTCAATCACAGTGATTTCTGA
 ATTGTGAGGGAAGTTATTAGCATGACAGGTGTCTGGTTCTGGCCCTGTACG
 ATTCCCATGAGTCAAGCAAATTGTAAGGGCTGGTCTATATCACACCCAAC
 CCAAGGATATGTCCCTCAAAGTCTAGCCCAGGCCCCGTCATCTTCAGC
 ATCATCTGGGAAACCAGGTCTGATTAGTAGTCCTTTAAGGAATACCTCTTA
 GGCTCCCATTTTACTGCTATCACAGAATCCAATAAAACCCTTACAGGAGAT
 TCAATGGGAAATGCTCAACACCCACTGTAGTTGGTGGTGACAATGACCAT

FIG. 2b

AATTTGGCTGTGCTGGATTTCAGGACAGAAAATTTGGGTGAAAGAGCAGGT
GAACAAAAGAGCTTCGACTTGCCCTAGCAGAGAGCAAGCCATACCATAACC
ACAAAGCCACAGCAATTACAACGGTGCAGTACCAGCACAGTAAATGAAC
AAAGTAGAGCCCAGAAACAGACCCAGAACTATATGAGGATTTAGTATACA
ATAAAGATGGTATTTTCGAGTCAGTAGGGAAAAGATGAATTATTCAATAAA
TGATGTTTGGCCAACTAGTAACCCATTTGGGAAAAAATAAAAAGTATGGTC
CCTACCTCACAGCATACACAAAAATAAATTCCAGACGGATTAAAATCTAA
ATGTAAAAAATAAAGCCATAAGTGGACTGGAAGAAAATAGAGAATTTTTT
TTAACATCCGTAGAAAGGGTAAAAACCCAGGCATGACATGAACCAAACT
GAAGAGGTTCTGTAACAAATACCCCCTTTTATATATTGGGCTCCAACAATA
AGAACCCATAGGAAAATGGAGAATGAACACAAATAGACAATTTATAGAA
GAGAAGGTTATAAGGTGTAAAATTATATCTATCTGAGAAACAAACACTAA
AACAAATGTGATTCTACTGTTCTCCACCCATACTGGCAAACTTAAGCCTG
ATAATATGCTGAGGGGAAATAAGCACTCTTGTTGGTGAGAGTATTAATTG
GCATAGCTTCTTTTGAAAATGACATAGCAATACCTGTTAAAATTGCAAAC
ATGCATGTCACTTAATCCAGTAATCCCCTTCTGGGAATCAATGCTACAAA
AACACTGACAAGTATACAAAGATACATTCAAGAGTGTTCACTGGGCCGGG
TGCGGTGGCTTCATGCCTGTAATCCCAGGGAGGCAGAGGCAAGACGATCG
CTTGACCCCAAGGAGTTCAAGGCCAGCCCGAGAAACACAGCAAGACCCTGT
CTCTCTTTTTTTTATTTAAAAAATAAATGTTCACTGTATCAGTTGTTCACAA
AAACAAACCAACATGTCCATTAACAGGGAACCATTTAATTAATCAAGTT
CATCTACACAATGTAATACCATGCAACTATTA AAAAGCACCTGATAATCC
AAAGCACACTGAGACAGAATAATGCTATTA AAAACACCAAGTAGTGGA
CACTGTGTTGCCTATGACACCATTTTTATTCAACATTTAAACAAATTTGTA
ACAGCAATTACATGAGTAGTGACAATGGCGTTTATGAGACTTTTCACTTTT
ATGTGCTTCTATTTTTGTTATGCTTCTATATACATCCATTTATTATGGAG
TGTTACTTTCAAAAATCACAAATGGGCCAGTATTATTTGGTGTTGCAAGGT
GAGCATATGACTTCTGATATCAACCTTTGCATATTACTTCTCAATTTAGGG
AAATTACAGACATCCCTTATTCTAACTAACTTAAAACCCAGCATTTCAAAC
ATACAGAATTGATGGGGAAAAAAAAGAAAGAAGAAAGAAAGAAAGGC
AACAAAGCTTCAGATGACAGTGACTCACATCAAATTATTTATAAAATCTGTT
AAATAGTGCCATCTTCTGGAGATACCTGGTATTACAGTCCAACCTCCAGTTG
ATGTCTTTACAGAGACAAGAGGAATAAAGGAAAAAATATTCAAGAACTG
AAAAGTATGGAGTCATGGAAAAATTGCTGTGATCCAAAGGCTACGGTGAT
AGGACAAGAAACAAGAGAACTCCAAGCAGTAAGACACTGCTGTTCTATTA
GCATCCAAACCTCCATACTCCTGTTTGCCCCAAGGCTTTTTTAAAAAATAG
AGACAGGATCTCACTATTTTGCTCAGGCTGGTCTTGAACCTCCTGGACTCAA
GCTATCCTCCTGCCTCGGCCTCCTAAAGTGCCGAGATTACAGGCTTGAGTC
ACCATACCTGGCTATTTATTTTTTCTTAACTCTCTTGCTGGCCTATAGCCA
CCATGGAAGCTAATAAAGAATATTAATTTAAGAGTAATGGTATAGTTCAC
TACATTGGAATACAGGTATAAGTGCCTACATTGTACATGAATGGCATAACA
TGGATCAATTACCCACCTGGGTGGCCAAAGGAAGTGCAGCAACCTCCCT
CCTTGGCTGTCTGGAACAAGCTTCCCACTAGATCCCTTTACTGAGTGCCTC
CCTCATCTTTAATTATGGTTAAGTCTAGGATAACAGGACTGGCAAAGGTG
AGGGGAAAGCTTCCTCCAGAGTTGCTCTACCTCTCCTCTACCGTCCTATC
TCCTCACTCCTCTCAGCCAAGGAGTCCAATCTGTCCTGAACTCAGAGCGTC
ACTGTCAACTACATAAAATTGCCAGAGAAGCTCTTTGGGACTACAAACAC
ATACCCTTAATGTCTTTATTTCTATTTTGTCTACCTCTTCAGTCTAGGTGAA
AAAATAGGAAGGATAATAGGGAAGAACTTTGTTTATGCCTACTTATCCGC
CCCTAGGAATTTTGAAAACCTCTAGGTAGCAATAAGAACTGCAGCATGGT

FIG. 2c

ATAGAAAAAGAGGAGGAAAGCTGTATAGAAATGCATAATAAATGGGCAG
 GAAAAGAAGCTTGCTTGGAACAAACAGGGAGGTTGAACTATAAGGAGAGAA
 AGCAGAGAGGCTAATCAACAAGGCTGGGTTCCCAAGAGGGCATGATGAG
 ACTATTACTAAGGTAGGAATTACTAAGGGCTCCATGTCCCCTTAGTGGCTT
 AGTACTATGTAGCTTGCTTTCTGCAGTGAACCTTCAGACCCTTCTTTTAGGA
 TCCTAGAATGGACTTTTTTTTTTTATCGGAAAACAGTCATTCTCTCAACATT
 CAAGCAGGCCCAAGTCTACCACACTCAATCACATTTTCTCTTCATATCAT
 AATCTCTCAACCATTCTCTGTCTTTTAACTGTTTTTCTATACCCTGATCAA
 ATGCCAACAAAAGTGAGAATGTTAGAATCATGTATTTTATAGAGGTAGACT
 GTATCTCAGATAAAAAAAAAAAGGGCAGATATTCCATTTTCCAAAATATGTA
 TGCAGAAAAAATAAGTATGAAAGGACATATGCTCAGGTAACAAGTTAATT
 TGTTTACTTGATTTTTATGAATTCCCTAAAACCTACGTCACCCGCCCGTTC
 CCACGCCCGCGCCACGTCACAACTCCACCCCTCATTATCATATTGGCT
 TCAATCCAAAATAAGGTATATTATTGATGATGTTAATTAACATGCATGGAT
 CCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAAATACCGCA
 TCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTT
 GGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCC
 ACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGC
 AAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGG
 CTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTG
 GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCT
 CCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCG
 CCTTTCTCCCTTCGGGAAGCGTGCGCTTTCTCATAGCTCACGCTGTAGGT
 ATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAAC
 CCCCCGTTAGCCCGACCGCTGCGCCTTATCCGGTAAGTATCGTCTTGAGT
 CCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC
 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTG
 GTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCT
 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCA
 AACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTA
 CGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGG
 TCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAG
 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGTTT
 TAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT
 GCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTATCCA
 TAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTA
 CCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGC
 TCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGA
 AGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGG
 GAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGC
 CATTGCTGCAGCCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTT
 CACGTAGAAAGCCAGTCCGCAGAAACGGTGCTGACCCCGGATGAATGTCA
 GCTACTGGGCTATCTGGACAAGGGAAAACGCAAGCGCAAAGAGAAAGCA
 GGTAGCTTGCAAGTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTTAT
 GGACAGCAAGCGAACCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTT
 GGGAAGCCCTGCAAAGTAACTGGATGGCTTTCTTGCCGCCAAGGATCTG
 ATGGCGCAGGGGATCAAGCTCTGATCAAGAGACAGGATGAGGATCGTTTC
 GCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTG
 GAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGA
 TGCCGCCGTGTTCCGGCTGTGAGCGCAGGGGCGCCCGGTTCTTTTGTCAA

FIG. 2d

GACCGACCTGTCCGGTGCCCTGAATGAACTGCAAGACGAGGCAGCGCGGC
TATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTG
TCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGGCGAAGTGCCGGGGCA
GGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGC
TGATGCAATGCGGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGA
CCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCC
GGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCC
AGCCGAACTGTTTCGCCAGGCTCAAGGCGAGCATGCCCCGACGGCGAGGATC
TCGTGCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAATAAT
GGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGC
TATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGG
CGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTC
GCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAATTTTGTTA
AAATTTTGTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAA
AATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTT
CAGTTTGGAACAAGAGTCCACTATTAAGAACGTGGACTCCAACGTCAAA
GGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC
CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACC
CTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGT
GGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCT
GGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTA
ATGCGCCGCTACAGGGCGCGTCCATTCGCCATTCAGGATCGAATTAATTCT
TAATTAA

FIG. 2e

SEQ ID NO:2 Human TM amino acid sequence

MLGVLVLGALALAGLGFPAPAEPPGGSQCV EHD CFALYPGPAT
FLNASQICDGLRGHLMTVRSSVAADVISLLLNGDGGVGRRRLWIGLQLPPGCGDPKR
LGPLRGFQWVTGDNNTSYSRWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQQ
CEVKADGFLCEFHFPA TCRPLAVEPGAAAAAVSITYGTPFAARGADFQALPVGSSAA
VAPLGLQLMCTAPPGAVQGHWAREAPGA WD C SVENGGCEHACNAIPGAPRCQCPA
GAALQADGRSCTASATQSCNDLCEHFCVPNPDQPGSYSCMCETGYRLAADQHRCED
VDDCILEPSPCPQRCVNTQGGFECHCYPNYDLVDGECVEPVDPCFRANCEYQCQPLN
QTSYLCVCAEGFAPIPHEPHRCQMFCNQTACPADCDPNTQASCECPEGYILDDGFICT
DIDECENGGFCSGVCHNLPGTFCICGPDSALARHIGTDCDSGKVDGGDSGSGEPPPS
PTPGSTLTPPAVGLVHSGLLIGISIASLCLV VALLALLCHLRKKQGAARAKMEYKCAA
PSKEVVLQHV RTERTPQRL

FIG. 3

SEQ ID NO:3 human TM nucleotide sequence

atgcttggg gtcctggcc ttggcgcgct ggccctggcc ggctgggggt tccccgcacc cgcagagccg cagccgggtg
gcagccagtg cgtcgagcac gactgcttcg cgctctaccc gggccccgcg accctcctca atgccagta gatctgcgac
ggactgcggg gccacctaata gacagtgcgc tctcggttg ctgccgatgt catttccttg ctactaacg gcgacggcgg
cgttggccgc cggcgccctt ggatcggcct gcagctgcca cccggctgcg gcgaccccaa gcgctcggg cccctgcgcg
gcttccagt gggtacggga gacaacaaca ccagctatag caggtgggca cggctcgacc tcaatggggc tccccctgc
ggccccgtgt gcgtcgctgt ctccgctgt gaggccactg tgcccagcga gccgatctgg gaggagcagc agtgcgaagt
gaaggccgat ggcttctct gcgagttcca ctccagcc acctgaggc cactggctgt ggagcccggc gccgcggctg
ccgccgtctc gatcacctac ggcaccccg tgcgggccc cggagcggac ttccaggcgc tgccggtggg cagctccgcc
gcggtggctc cctcggctt acagctaag tgcaccgcgc cggcgggagc ggtccagggg cactgggcca gggaggcgcc
gggcgcttg gactgcagc tgagaaacgg cggctgcgag caccgctgca atgcgatccc tggggctccc cgtgccagt
gcccagccgg cggcgccctg caggcagacg ggcgctcctg caccgcatcc gcgacgcagt cctgcaacga cctctgcgag
cacttctgcg tccccaaccc cgaccagccg ggctcctact cgtgcatgtg cgagaccggc taccggtgg cggccgacca
acaccggtgc gaggacgtgg atgactgcat actggagccc agtccgtgtc cgcagcgtg tgtcaacaca cagggtggct
tcgagtcca ctgtaccct aactacgacc tggtagcgg cgagtgtgtg gagccgtgg acccgtgctt cagagccaac
tgcgagtacc agtgccagcc cctgaacaa actagctacc tctgcgtctg cggcagggc ttgcgcca tccccacga
gccgcacagg tgccagatgt tttgaacca gactgcctgt ccagccgact gcgacccaa caccaggct agctgtgagt
gccctgaagg ctacatcctg gacgacggtt tcatctgcac ggacatcgac gagtgcgaaa acggcggtt ctgtccggg
gtgtgccaca acctccccgg tacctcgag tgcctctgcg ggcccgactc ggcccttgc cggcacattg gcaccgactg
tgactccggc aaggtggacg gtggcgacag cggctctggc gagccccgc ccagcccgac gcccggtcc acctgactc
ctccggccgt ggggctcgtg cattcgggt tgcctatagg catctccatc gcgagcctgt gcctggtggt ggcgctttg
gcgctcctt gccacctgc caagaagcag ggcgccgcca gggccaagat ggagtacaag tgcgcggccc ctccaagga
ggtagtgct cagcacgtgc ggaccgagcg gacgccgac agactc

SEQ ID NO: 4

GTTTAAACGGGCCCCTCTAGACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTAC
GGGGTCATTAGTTCATAGCCCATGATATCATATGGAGTTCCGCGTTACATAACTTACGGTAAATGG
CCCGCCTGGCTGACCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGT
AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCACTTGCC
AGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCG
CCTGGCATTATGCCAGTNCATGACCTTATGGGACTTTCTACTTGGCAGACATCTACGTATTAGTC
ATCGCTATTACCATGGTGATGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCA
CGGGGATTTTCCAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG
GGACTTTTCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGT
GGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCTGCTTACTGGCTTATCGAGAT
ATCTGCAGAATTTCATCTGTCTGACTGCTACCGGCAGCGCGCAAGAAGTGTCTGGGCTGGG
ACGGACAGGAGAGGCTGTGCGCATCGCGTCTGTGCCCTCTGCTCCGGCACGCAATGTCTGCGAG
TGCCCGCGCTTTCCCGGCGCCTGCACGCGGCGCGCCTGGGTAACATGCTTGGGGTCTGGTCTCT
GGCGCGCTGGCCCTGGCCGGCCTGGGGTTCCCCGCACCCGCAGAGCCGCAGCCGGGTGGCAGCCA
GTGCGTCGAGCACGACTGCTTCGCGCTCTACCCGGGCCCCGCGACCTTCCTCAATGCCAGTCAGAT
CTGCGACGGACTGCGGGGCCACCTAATGACAGTGCCTCTCGGTGGCTGCCGATGTCATTTCTCT
GCTACTGAACGGCGACGGCGGCGTTGGCCGCCGGCGCCTCTGGATCGGCCTGCAGCTGCCACCCG
GCTGCGGCGACCCCAAGCGCCTCGGGCCCCCTGCGCGGCTTCCAGTGGGTTACGGGAGACAACAAC
ACCAGCTATAGCAGGTGGGCACGGCTCGACCTCAATGGGGCTCCCCTCTGCGGCCCGTTGTGCGTC
GCTGTCTCCGCTGCTGAGGCCACTGTGCCAGCGAGCCGATCTGGGAGGAGCAGCAGTGCGAAGT
GAAGGCCGATGGCTTCTCTGCGAGTTCCACTTCCAGCCACCTGCAGGCCACTGGCTGTGGAGCC
CGGCGCCGCGGCTGCCGCGTCTCGATCACCTACGGCACCCCGTTTCGCGGCCCGCGGAGCGGACTT
CCAGGCGCTGCCGCTGGGCAGCTCCGCGCGGCTGGCTCCCCTCGGCTTACAGCTAATGTGCACCGC
GCCGCCCGGAGCGGTCCAGGGGCACTGGGCCAGGGAGGCGCCGGGCGCTTGGGACTGCAGCGTG
GAGAACGGCGGCTGCGAGCACGCGTGCAATGCGATCCCTGGGGCTCCCCGCTGCCAGTGCCACAGC
CGGCGCCGCCCTGCAGGCAGACGGGCGCTCTGCACCGCATCCGCGACGCAGTCCTGCAACGACC
TCTGCGAGCACTTCTGCGTTCCCAACCCCGACCGCCGGGCTCCTACTCGTGCATGTGCGAGACCG
GCTACCGGCTGGCGGCCGACCAACACCGGTGCGAGGACGTGGATGACTGCATACTGGAGCCCCAGT
CCGTGTCCGCAGCGCTGTGTCAACACACAGGGTGGCTTCGAGTGCCACTGCTACCCTAACTACGAC
CTGGTGGACGGCGAGTGTGTGGAGCCCCGTGGACCCGTGCTTCAGAGCCAACTGCGAGTACCAGTG
CCAGCCCCGTAACCAAACTAGCTACCTCTGCGTCTGCGCCGAGGGCTTCGCGCCCATTCCCCACGA
CCGCGACAGGTGCCAGATGTTTTGCAACCAAGACTGCGCTGTCCAGCCGACTGCGACCCCAACCCCA
GGCTAGCTGTGAGTGCCCTGAAGGCTACATCCTGGACGACGGTTTCATCTGCACGGACATCGACGA
GTGCGAAAACGGCGGCTTCTGCTCCGGGGTGTGCCACAACCTCCCCGGTACCTTCGAGTGCATCTG
CGGGCCCGACTCGGCCCTTGCCCGCCACATTGGCACCGACTGTGACTCCGGCAAGGTGGACGGTG
GCGACAGCGGCTCTGGCGAGCCCCCGCCAGCCCGACGCCCGGCTCCACCTTGACTCCTCCGGCCG
TGGGGCTCGTGCATTGCGGCTTGCTCATAGGCATCTCCATCGCGAGCCTGTGCCTGGTGGTGGCGC
TTTTGGCGCTCCTCTGCCACCTGCGCAAGAAGCAGGGCGCCGCCAGGGCCAAGATGGAGTACAAG
TGCGCGGCCCTTCCAAGGAGGTAGTGCTGCAGCACGTGCGGACCGAGCGGACGCCGAGAGACT
CTGAGCGGCCTCCGTCCAGGAGCCTGGCTCCGTCCAGGAGCCTGTGCCTCCTCACCCCCAGCTTTG
CTACCAAAAGCACCTTAGCTGGCATTACAGCTGGAGAAGACCTCCCCGCACCCCCCAAGCTGTTTT
CTTCTATTCCATGGCTAACTGGCGAGGGGTGATTAGAGGGAGGAGAATGAGCCTCGGCCTCTTCC
GTGACGTCACTGGACCACTGGGCAATGATGGCAATTTTGTAAACGAAGACACAGACTGCGATTTGTC
CCAGGTCCTCACTACCGGGCGCAGGAGGGTGAGCGTTATTGGTCGGCAGCCTTCTGGGCAGACCTT
GACCTCGTGGGCTAGGGATGACTAAAATATTTATTTTTTTAAGTATTTAGGTTTTTGTGTTTCTCT
TTGTTCTTACCTGTATGTCTCCAGTATCCACTTTGCACAGCTCTCCGGTCTCTCTCTCTACAACT
CCCACTTGTCATGTGACAGGTAACTATCTTGGTGAATTTTTTTTCTAGCCCTCTCACATTTATG
AAGCAAGCCCCACTTATCCCCATTCTTCTAGTTTTCTCTCTCCAGGAACTGGGCCAACTCACCTG
AGTCACCCTACCTGTGCCTGACCCTACTTCTTTTGTCTTAGCTGTCTGCTCAGACAGAACCCTAC
ATGAAAACAGAAACAAAAACACTAAAAATAAAAAATGGCCATTTGCTTTTTTACCAGATTTGCTAATT
TATCCTGAAATTTAGATTCCCAGAGCAAAATAATTTTAAACAAAGGTTGAGATGTAAAAGGTATT
AAATTGATGTGCTGGACTGTATGAAATTAACCCAAAGAGGTATTTATCTTTACTTTTAAACA
GTGAGCCTGAATTTTGTGCTGTTTTGATTGTACTGAAAAATGGTAATTGTTGTAATCTTCTTAT
GCAATTTCTTTTTTGTATTATTACTTATTTTTGACAGTGTTGAAAATGTTTCAAGAGGTTGCTCTAG
ATTGAGAGAAGAGACAAACACCTCCCAGGAGACAGTTCAAGAAAGCTTCAAACCTGCATGATTCAT
GCCAATTAGCAATTGACTGTCACTGTTCTTGTCACTGGTAGACCAAAATAAAACCAGCTCTACTG
GTCTTGTTGGAATTGGGAGCTTGGGAATGGATCCTGGAGGATGCCCAATTAGGGCCTAGCCTTAATC

FIG. 5

AGGTCCTCAGAGAATTTCTACCATTTTCAGAGAGGCCTTTTGGAATGTGGCCCCTGAACAAGAATTG
GAAGCTGCCCTGCCCATGGGAGCTGGTTAGAAATGCAGAATCCTAGGCTCCACCCCATCCAGTTCA
TGAGAATCTATATTTAACAAGATCTGCAGGGGGTGTGTCTGCTCAGTAATTTGAGGACAACCATT
CAGACTGCTTCCAATTTTCTGGAATACATGAAATATAGATCAGTTATAAGTAGCAGGCCAAGTCAG
GCCCTTATTTTCAAGAACTGAGGAATTTTCTTTGTGTAGCTTTGCTCTTTGGTAGAAAAGGCTAGG
TACACAGCTCTAGACACTGCCACACAGGGTCTGCAAGGTCTTTGGTTCAGCTAAGCTAGGAATGAA
ATCCTGCTTCAGTGTATGGAAATAAATGTATCATAGAAATGTAACCTTTTGTAAGACAAAGGTTTT
CTCTTCTATTTTGTAACCTCAAAATATTTGTACATAGTTATTTATTTATTGGAGATAATCTAGAACA
CAGGCAAAATCCTTGCTTATGACATCACTTGTACAAAATAAACAAATAACAATGTGAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGTAGCAGTCGACAGATGAATTCCACCACACTG
GACTAGTGGATCCGAGCTCGGTACCAAGCTTAAGTTTAAAC

FIG. 5a

SEQ ID NO 5

TCTAGACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCA
TAGCCCATGATATCATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACC
GCCCAACGACCCCCGCCCATGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAC
TTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTA
TCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCC
AGTNCATGACCTTATGGGACTTTCCTACTTGGCAGACATCTACGTATTAGTCATCGCTATTACCATG
GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTTCCAAG
TCTCCACCCCATTTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG
TCGTAACAACCTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTCTCTGGCTAACTAGAGAACCCCTGCTTACTGGCTTATCGAGATATC

FIG. 6

SEQ ID NO 6

GGCAGCGCGCAGCGGCAAGAAGTGTCTGGGCTGGGACGGACAGGAGAGGGCTGTCGCCATCGGCG
TCCTGTGCCCCCTCTGCTCCGGCACGGCCCTGTGCGAGTGCCCGCGCTTTCCCCGGCGCCTGCACGC
GGCGCGCCTGGGTAACATGCTTGGGGTCTGTGTCCTTGGCGCGCTGGCCCTGGCCGGCCTGGGGTT
CCCCGACCCGCGAGAGCCGCGAGCCGGGTGGCAGCCAGTGCCTCGAGCACGACTGCTTCGCGCTCT
ACCCGGGCCCCGCGACCTTCCTCAATGCCAGTCAGATCTGCGACGGAAGTGGGGGCCACCTAATG
ACAGTGCCTCTCGGTGGCTGCCGATGTCAATTTCTTGTCTACTGAACGGCGACGGCGGCGTTGGC
CGCCGGCGCCTCTGGATCGGCCTGCAGCTGCCACCCGGCTGCGGCGACCCCAAGCGCCTCGGGCC
CCTGCGCGGCTTCCAGTGGGTACGGGAGACAACAACACCAGCTATAGCAGGTGGGCACGGCTCG
ACCTCAATGGGGCTCCCTCTGCGGCCCGTTGTGCGTCGCTGTCTCCGCTGCTGAGGCCACTGTGC
CCAGCGAGCCGATCTGGGAGGAGCAGCAGTGCAGAGTGAAGGCCGATGGCTTCCTCTGCGAGTTC
CACTTCCCAGCCACCTGCAGGCCACTGGCTGTGGAGCCCCGGCGCCGGCTGCCGCCGTCTCGATC
ACCTACGGCACCCCGTTTCGCGGCCCGCGGAGCGGACTTCCAGGCGCTGCCGGTGGGCAGCTCCGC
CGCGGTGGCTCCCCCTCGGCTTACAGTAACTGTCACCGCGCCCGCGGAGCGGCTCCAGGGGCACT
GGGCGAGGAGGGCGCGGCGCTTGGGACTGCAGCGTGGAGAACGGCGGCTGCGAGCACGCGTG
CAATGCGATCCCTGGGGCTCCCCGCTGCCAGTGCCAGCCGGCGCCGCCCTGCAGGCAGACGGGC
GCTCCTGCACCGCATCCGCGACGCGAGTCTGCAACGACCTCTGCGAGCACTTCTGCGTTCCCAACC
CCGACCAGCCGGGCTCCTACTCGTGATGTGCGAGACCGGCTACCGGCTGGCGGCCGACCAACAC
CGGTGCGAGGACGTGGATGACTGCATACTGGAGCCCAGTCCGTGTCCGCAGCGCTGTGTCAACAC
ACAGGGTGGCTTCGAGTGCCACTGCTACCCTAAGTACGACCTGGTGGACGGCGAGTGTGTGGAGC
CCGTGGACCCGTGCTTCAGAGCCAAGTGCAGTACCAGTGCAGCCCTGAACCAAAGTACGCTAC
CTCTGCGTCTGCGCCGAGGGCTTCGCGCCCATTCACACGAGCCGCACAGGTGCCAGATGTTTTGC
AACCAGACTGCCTGTCCAGCCGACTGCGACCCCAACACCCAGGCTAGCTGTGAGTGCCCTGAAGG
CTACATCTGGACGACGGTTTCATCTGACGAGACATCGACGAGTGCGAAAACGGCGGCTTCTGCTC
CGGGTGTGCCACAACCTCCCGGTACCTTCGAGTGCATCTGCGGGCCCGACTCGGCCCTTGCCCG
CCACATTGGCACCGACTGTGACTCCGGCAAGGTGGACGGTGGCGACAGCGGCTCTGGCGAGCCCC
CGCCAGCCCGACGCCCCGGCTCCACCTTGACTCCTCCGGCCGTGGGGCTCGTGCAATTCGGGCTTGC
TCATAGGCATCTCCATCGCGAGCCTGTGCTGCTGGTGGTGGCGCTTTTGGCGCTCCTCTGCCACCTGCG
CAAGAAGCAGGGCGCCGCCAGGGCCAAGATGGAGTACAAGTGCAGCGGCCCTTCCAAGGAGGTA
GTGCTGCAGCACGTGCGGACCGAGCGGACCGCGAGAGACTCTGAGCGGCCCTCCGTCCAGGAGCC
TGGCTCCGTCCAGGAGCCTGTGCTCCTCACCCCCAGCTTTGCTACCAAAGCACCTTAGCTGGCAT
TACAGCTGGAGAAGACCCCTCCCCGCACCCCCCAAGCTGTTTTCTTCTATTCCATGGCTAACTGGCG
AGGGGGTGATTAGAGGGAGGAGAATGAGCCTCGGCCTCTTCCGTGACGTCACTGGACCACTGGGC
AATGATGGCAATTTTGTAAACGAAGACAGACAGTGGCATTTGTCCAGGTCTCACTACCGGGCGCA
GGAGGGTGAGCGTTATTGGTCCGCGAGCCTTCTGGCGAGACCTTGACCTCGTGGGCTAGGGATGACT
AAAATATTTATTTTAAAGTATTTAGGTTTTTGTGTTTCTTTGTTTCTTACCTGTATGTCTCCAG
TATCCACTTTGCACAGCTCTCCGGTCTCTCTCTCTACAACTCCCACTTGTCATGTGACAGGTAA
ACTATCTTGGTGAATTTTTTTTCTAGCCCTCTCACATTTATGAAGCAAGCCCCACTTATTTCCCAT
TCTTCTAGTTTTCTCTCCAGGAAGTGGGCAACTCACCTGAGTCACCCTACCTGTGCTGACCC
TACTTCTTTGCTCTTAGCTGTCTGCTCAGACAGAACCCCTACATGAAACAGAAACAAAAACACTA
AAAATAAAATGGCCATTTGCTTTTTACACAGATTTGCTAATTTATCCTGAAATTTAGATTCCCAG
AGCAAAATAATTTTAAACAAAGGTTGAGATGTAAAAGGTATTAATTTGATGTTGCTGGACTGTCAT
AGAAATTACACCCAAAGAGGTATTTATCTTTACTTTTAAACAGTGAGCCTGAATTTTGTGCTGTTT
TGATTTGTACTGAAAAATGGTAATTTGCTAATCTTCTTATGCAATTTCTTTTTTGTATTATTAC
TTATTTTGTACAGTGTTGAAAATGTTTCAAGAGGTTGCTCTAGATTGAGAGAAGAGACAAACACCTC
CCAGGAGACAGTTCAAGAAAGCTTCAAACTGCATGATTGATGCAATTTAGCAATTTGACTGTCACTG
TTCTTGTCACTGGTAGACCAAAATAAAACCAGCTCTACTGGTCTTGTGGAATTTGGGAGCTTGGGA
ATGGATCCTGGAGGATGCCCAATTAGGGCCTAGCCTTAATCAGGTCTCAGAGAATTTCTACCATT
TCAGAGAGGCCTTTTGAATGTGGCCCTGAACAAGAATTGGAAGCTGCCCTGCCCATGGGAGCT
GGTTAGAAATGCAGAATCTAGGCTCCACCCATCCAGTTCATGAGAATCTATATTTAACAAGATC
TGCAGGGGGTGTGTCTGCTCAGTAATTTGAGGACAACCATTCAGACTGCTTCCAATTTTCTGGAA
TACATGAAATATAGATCAGTTATAAGTAGCAGGCCAAGTCAGGCCCTTATTTTCAAGAACTGAG
GAATTTTCTTTGTGTAGCTTTGCTCTTTGGTAGAAAAGGCTAGGTACACAGCTCTAGACACTGCCA
CACAGGGTCTGCAAGGTCTTTGGTTAGCTAAGCTAGGAATGAAATCCTGCTTCAGTGTATGGAAA
TAAATGTATCATAGAAATGTAACTTTTGTAAGACAAAGGTTTTCTCTTCTATTTTGTAACTCAAA
ATATTTGTACATAGTTATTTATTTATTGGAGATAATCTAGAACACAGGCCAAATCCTTGCTTATGAC

FIG. 7

ATCACTTGTACAAAATAAACAAATAACAATGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAA

FIG. 7a